

Visualización de información en 2016: conceptos, contenidos y tecnología

Information visualization in 2016: concepts, content and technology

Mario Pérez-Montoro

Pérez-Montoro, Mario (2017). "Visualización de información en 2016: conceptos, contenidos y tecnología". *Anuario ThinkEPI*, v. 11, pp. 284-296.

<https://doi.org/10.3145/thinkepi.2017.55>

Publicado en *IweTel* el 21 de diciembre de 2016



Resumen: La visualización de información se caracteriza por cubrir las funciones básicas de almacenamiento y comunicación de datos y de soporte al razonamiento. Estas funciones se encuentran detrás de la mayoría de las visualizaciones desarrolladas en los últimos tiempos. Sin embargo, entre éstas, destaca un pequeño subconjunto por sus planteamientos innovadores y su contribución al desarrollo de la propia disciplina de la visualización. El objetivo de este trabajo es revisar algunas de esas visualizaciones innovadoras desarrolladas a lo largo de 2016. Para abordarlas las hemos organizado en tres grupos. Por un lado, presentaremos una selección de aquellas que han destacado por incorporar novedades conceptuales y nuevas propuestas de visualización. A continuación, mostraremos otras que han llamado la atención por la importancia del conjunto de datos que han visualizado. Por último,

daremos cuenta de algunos de los avances tecnológicos que se han desarrollado dentro de la disciplina a lo largo de este año.

Palabras clave: Visualización de la información; 2016; Tendencias; Visualización de datos; Visualización interactiva; Infografía; Visualizaciones multimedia; Conceptos; Contenidos; Tecnología.

Abstract: Information visualization covers the basic functions of data storage and communication, and reasoning support; most of the recently developed visualizations include these functions. However, among these, a small subset stands out for their innovative approaches and contribution to the development of the discipline. This article will review some of the innovative visualizations developed in 2016 and organize them into three groups. On the one hand, we present a small selection of visualizations that incorporate conceptual novelties and new visual proposals. On the other hand, we show some others that have drawn attention to the importance of the data set they have visualized. Finally, we account for some of the technological advances that have been developed within the discipline throughout this year.

Keywords: Information visualization; 2016; Trends; Data visualization; Interactive visualization; Infographics; Multimedia visualization; Concepts; Contents; Technology.

1. Introducción

Existe cierto consenso a la hora de identificar las tres principales funciones básicas de la disciplina de la visualización de información:

- la utilización de esta disciplina nos ofrece una forma cómoda de almacenamiento de información: mediante una visualización podemos registrar en una representación gráfica una importante cantidad de datos;
- facilita la comunicación efectiva de información: esas representaciones gráficas permiten transmitir datos de manera que puedan ser asimilados de forma cómoda por los receptores a los que van dirigidos;
- actúa de soporte al razonamiento: graficar un conjunto de datos permite aflorar e identificar, mediante representaciones visuales, aspectos, relaciones y patrones que no son evidentes o perceptibles abordando a simple vista ese con-

junto de datos. Ese efecto se consigue porque transformamos o representamos entidades categoriales y numéricas (entidades abstractas, en definitiva) utilizando objetos (puntos, barras o líneas, principalmente) dotados de propiedades espaciales y físicas (longitud, anchura, ubicación o color, entre otras). En este sentido, logramos pasar de lo abstracto (que nos resulta difícil de computar) a lo físico, mucho más sencillo de operar y computar para nuestro sistema cognitivo.

“Graficar un conjunto de datos permite aflorar e identificar, mediante representaciones visuales, aspectos, relaciones y patrones que no son evidentes o perceptibles abordando a simple vista ese conjunto de datos”

Estas tres funciones (y, de forma especial, la del soporte al razonamiento) se encuentran detrás de la mayoría de las visualizaciones desarrolladas en los últimos tiempos. Sin embargo, entre estas, destacan un pequeño subconjunto por sus planteamientos innovadores y su contribución al desarrollo de la propia disciplina.

El objetivo de este trabajo es revisar algunas de esas visualizaciones innovadoras llevadas a cabo a lo largo de 2016. Para abordarlas las hemos

organizado en tres grupos:

- selección de aquellas que han destacado por incorporar novedades conceptuales y nuevas propuestas de visualización;
- otras que han llamado la atención por la importancia del conjunto de datos que han visualizado;
- algunos de los avances tecnológicos que se han llevado a cabo dentro de la disciplina a lo largo de este año.

2. Conceptos

El primer grupo de visualizaciones que queremos destacar fundamentan su singularidad en la aportación de nuevas propuestas visuales para la representación de datos. Más allá de las clásicas gráficas de barras (o columnas), de líneas o de puntos, aportan nuevas fórmulas gráficas para transformar datos en representaciones visuales nunca antes propuestas.

Dentro de este grupo, en 2016, destacan visualizaciones relacionadas con:

- figuras originales dotadas de nuevas capacidades de representación;
- codificación de los diferentes aspectos musicales de una pieza musical;
- reinterpretación interactiva del clásico *tree-map*;
- propuesta visual botánica, arbórea, para la representación de textos y su estructura jerárquica.

Register for free at <https://www.scipedia.com> to download the version without the watermark

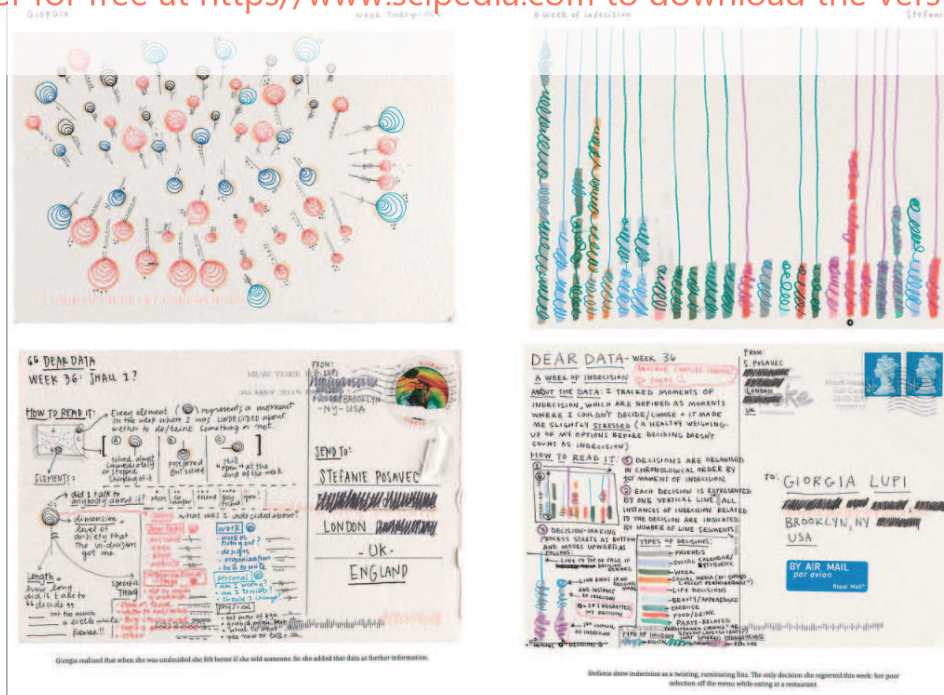


Figura 1. Visualizaciones (junto a sus leyendas) alternativas de los momentos de indecisión sufridos por ambas diseñadoras durante una semana enviadas por correo ordinario en forma de postales
<http://www.dear-data.com>

La primera propuesta de interacción personal entre dos diseñadoras, Giorgia Lupi y Stefanie Posavec; la primera ubicada en la ciudad de New York y la segunda en Londres. Estas dos profesionales, una vez por semana y durante un año completo, pactaron un listado de tipos de datos relacionados con sus vidas personales, los recolectaron y visualizaron de forma independiente y los resultados se los intercambiaron recíprocamente enviándoselos a través de postales por correo ordinario. En el anverso de cada postal dibujaron a mano la visualización semanal; en el reverso

so, la leyenda o instrucciones para interpretar la visualización de forma correcta (figura 1). Se visualizaron durante el período de una semana, entre otras muchas cosas curiosas, las puertas que abrieron, las veces y las causas de sus quejas, sus momentos de indecisión, sus risas, sus despedidas, o las veces y los aparatos que consultaron para saber qué hora era. Las visualizaciones resultantes se alejaban de las propuestas clásicas como las gráficas de puntos, barras o líneas; combinando nuevas figuras con diferentes capacidades de representación.

El proyecto se recoge en el sitio web *Dear Data* (Lupi; Posavec, 2016), ha sido expuesto en diversos museos (como el *London's Science Museum*, entre otros), ha ganado diversos premios y nominaciones en el campo del diseño (como el de *Most Beautiful* en *The Information is Beautiful Awards* o el *The Design Museum's illustrious '2016 Designs of the Year'*) y ha sido publicado en forma de libro en Estados Unidos por la editorial Princeton Architectural Press y en Inglaterra por Penguin. Es importante señalar también que en noviembre de 2016 el proyecto fue adquirido por el *Museum of Modern Art (MoMA)* de Nueva York como parte de su colección permanente. <http://www.dear-data.com>

"Transformamos o representamos entidades categoriales y numéricas (entidades abstractas, en definitiva) utilizando objetos (puntos, barras o líneas, principalmente) dotados de propiedades espaciales y físicas (longitud, anchura, ubicación o color, entre otras)"

Otro proyecto innovador de 2016 ha sido el de la *Toronto Symphony Orchestra (TSO, 2016)*. Hannah Chan-Hartley, editora y musicóloga de la TSO, junto al diseñador gráfico Gareth Fowler y la participación de la agencia *Haft2*, crearon unas guías visuales para facilitar a sus espectadores la audición de las obras que ejecutan. Frente a otras alternativas exclusivamente textuales, estas guías visuales, incluidas en los libros de mano, incorporan también material gráfico que permite visualizar la progresión temática de la música mostrando los instrumentos ejecutados y, mediante notación basada en código morse, la duración de los mismos.

De forma más concreta, cada movimiento sinfónico de la obra se representa en una única página para facilitar su lectura. En esta página, los distintos puntos y guiones corresponden a

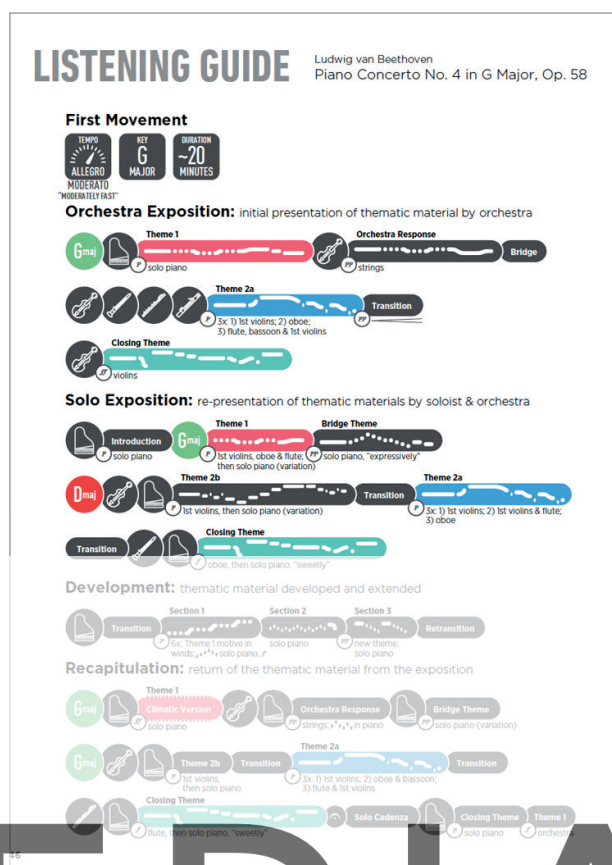


Figura 2. Guía visual para la audición del *Concerto para piano y orquesta número 4 en sol mayor, Opus 58* de Beethoven

<https://goo.gl/9yJOWt>

Register for free at <https://www.scipedia.com> to download the version without the watermark

los valores de duración en la notación musical estándar y siguen el mismo patrón que se puede observar en la partitura. Además de los puntos y guiones, se utilizan otras formas para representar otros gestos musicales ("curvas" y "olas" para motivos ascendentes y descendentes, entre otros). Los temas musicales principales se representan asignándoles un color. La clave en la que se centra un movimiento particular de la sinfonía también se representa mediante el color. La elección de los colores utilizados para los temas y las claves intentan representar el carácter del elemento y también el estado de ánimo del movimiento. Los temas que aparecen en múltiples movimientos están en el mismo color. La visualización se completa con un conjunto de iconos para representar los diversos instrumentos de la orquesta y se acompaña con una leyenda explicativa de los diversos símbolos incluidos (figura 2).

La tercera visualización que queremos destacar ha sido publicada por el medio digital *Five Thirty Eight* (Casselman; Conlen; Fischer-Baum, 2016).

<http://fivethirtyeight.com>

Es una aplicación original e interactiva del clásico *TreeMap* (o mapa de árbol). Creado por Ben Shneiderman (*University of Maryland*) en la década de 1990 para representar de forma concentrada carpetas, subcarpetas y ficheros en un disco duro de un ordenador, un *treemap* es una propuesta de visualización consistente en la disposición de superficies rectangulares anidadas jerárquicamente en un rectángulo mayor y se utiliza para visualizar en el mínimo espacio posible grandes conjuntos de datos que se encuentran estructurados de una forma categorial o jerárquica. En el caso que nos ocupa, se representa en un gran rectángulo el total anual (media de los últimos 10 años) de muertes por disparo de arma de fuego en Estados Unidos (figura 3). Los datos han sido extraídos de la base de datos *Multiple cause of death* de los *Centers for Disease Control and Prevention*. A su vez, ese gran rectángulo está formado por la agregación de más de 33.000 pequeños cuadrados que representan, cada uno, una muerte por disparo. Mediante un sistema de interacción simple y limpio basado en una navegación secuencial, el visitante de la página puede asistir a una narración donde explican las cifras de esas muertes por arma de fuego. En este sentido nos muestra que, aunque tendamos a fijarnos en las muertes derivadas del terrorismo, tiroteos, policías muertos en el cumplimiento del deber y disparos a civiles por parte de los cuerpos de seguridad, casi dos tercios de las muertes por armas de fuego son suicidios. Más del 85 por ciento de las víctimas del suicidio son hombres, y más de la mitad de todos los suicidios son hombres de 45 años o más.

Otro tercio de todas las muertes por armas de fuego (alrededor de 12.000 en total cada año) son homicidios. Más de la mitad de las víctimas de homicidio son hombres jóvenes, dos tercios de los cuales son afroamericanos. Las mujeres son mucho menos propensas a ser víctimas de homicidios con armas de fuego (alrededor de 1.700 de ellas mueren cada año, muchas en incidentes de violencia doméstica). Las restantes muertes por armas de fuego son accidentes o se clasifican como indeterminadas. La visualización desemboca en una pantalla final en la que se puede explorar de forma interactiva la totalidad del *treemap* mediante el filtrado por causa, género, edad y raza.

“Stefanie Posavec (proyecto *Dear Data*) y Greg McInerny (*Microsoft Research*) han creado una innovadora manera de representar y comparar unidades textuales utilizando para ello formas visuales botánicas, arbóreas, como si de un organismo vegetal se tratara”

La última de las visualizaciones que queremos destacar gracias a su aportación de nuevas propuestas visuales se centra en el tema de la representación textual, aspecto crítico y central en la nueva disciplina de las humanidades digitales. Concretamente, **Stefanie Posavec** (2016) (cocreadora del proyecto

Register for free at <https://www.scipedia.com> to download the version without the watermark



Figura 3. *Treemap* interactivo para visualizar las muertes por disparo de arma de fuego en Estados Unidos <http://fivethirtyeight.com/features/gun-deaths>

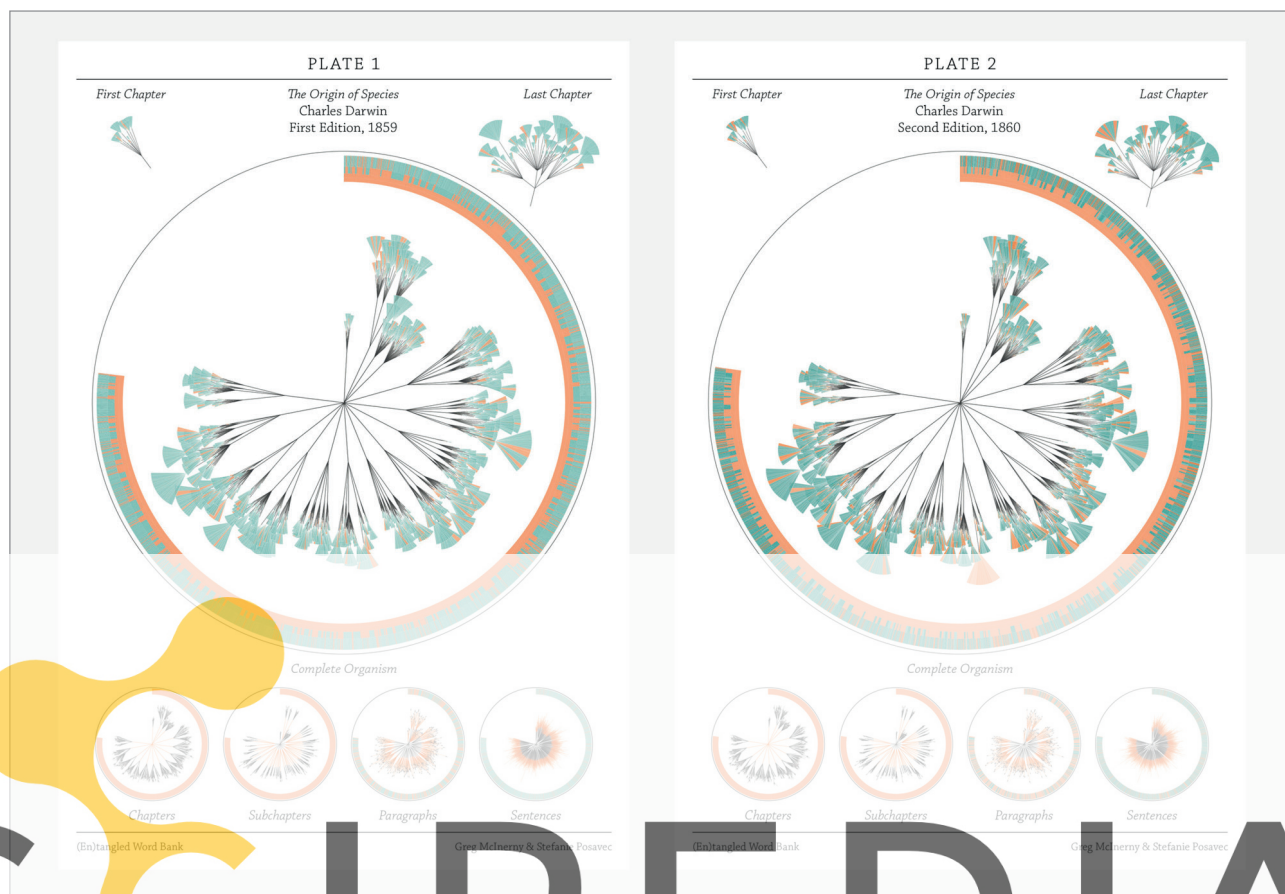


Figura 4. Visualización botánica del texto de la primera y segunda edición de la obra *El origen de las especies* del científico Charles Darwin
<https://goo.gl/638Az5>

Register for free at <https://www.scipedia.com> to download the version without the watermark

Microsoft Research) han creado una innovadora manera de representar y comparar unidades textuales utilizando para ello formas visuales botánicas, arbóreas, como si de un organismo vegetal se tratara. En este proyecto, denominado *(En)tangled Word Bank*, como caso donde aplicar su propuesta eligen las seis primeras ediciones de la obra *El origen de las especies* del científico Charles Darwin publicadas a lo largo de su vida.

Utilizando esos textos, crean una rueda que representa cada edición del libro (figura 4). En el interior de la rueda, se muestran de forma jerárquica y mediante una estructura en forma de ramas y hojas (con variaciones de tamaño), los capítulos, sus respectivos subcapítulos o secciones, los párrafos y los enunciados que los conforman. Los enunciados se representan coloreados de azul si vuelven a aparecer en la siguiente edición o de naranja si no lo hacen. Esta variación de colores permite aflorar y revelar los avances científicos y los ajustes que aporta cada nueva edición, además de cómo se resuelven las secciones que entran en conflicto. Esa misma estructura representativa se repite en el círculo exterior de cada edición.

3. Contenidos

El segundo grupo de visualizaciones que queremos recoger destaca por haber sido capaces de representar conjuntos de datos que se caracterizan por su originalidad o por el interés despertado a lo largo de 2016. Entre éstas destacan algunas visualizaciones dedicadas a:

- representar datos relacionados con dos de los eventos más importantes de este año;
- mostrar los contrastes existentes entre diferentes culturas;
- abordar el nuevo campo de las humanidades digitales.

Los juegos olímpicos de Río de Janeiro y las elecciones presidenciales norteamericanas han sido dos de los hechos más destacados de 2016. Estos eventos, por sus características propias, han generado infinidad de datos. La mayoría de los medios de comunicación han creado diferentes propuestas visuales para completar sus textos y poder comunicar de forma efectiva a sus lectores y espectadores los aspectos relevantes contenidos en esos conjuntos de datos.

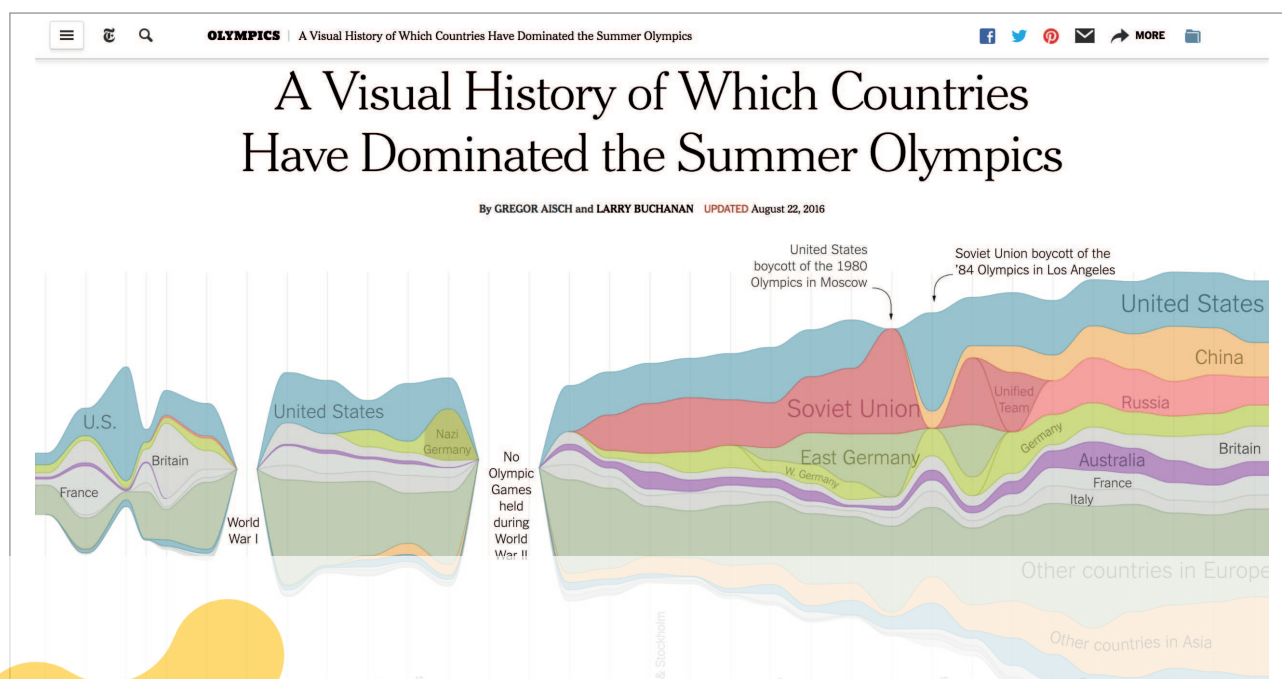


Figura 5. Historia del dominio de algunos países en los juegos olímpicos de verano

http://www.nytimes.com/interactive/2016/08/08/sports/olympics/history-olympic-dominance-charts.html?_r=1

Respecto a los juegos olímpicos de Río, destacamos tres visualizaciones que nos han parecido especialmente interesantes.

“Los juegos olímpicos de Río de Janeiro y las elecciones presidenciales norteamericanas, dos de los hechos más destacados de 2016, han generado infinidad de datos y propuestas visuales”

Register for free at <https://www.scipedia.com> to download the version without the watermark

Por un lado, el excelente trabajo realizado por el departamento de gráficos del diario norteamericano *The New York times*. A lo largo del evento olímpico, el diario neoyorquino ha ofrecido diversas visualizaciones interesantes. Pero destaca una en la que, aprovechando el hecho noticioso de los juegos y utilizando gráficas de áreas apiladas, ofrecen un análisis visual (global o agregado y por deporte) de la historia del dominio de algunos países en los juegos olímpicos de verano (Aisch; Buchanan, 2016) (figura 5). Destaca observar el dominio de la antigua URSS durante su existencia y como algunos países destacan en algún deporte concreto (como, por ejemplo, Etiopía en maratón; China en tenis de mesa; o Italia en esgrima) por encima de las potencias atléticas tradicionales.

Por otro lado, queremos destacar el interesante trabajo realizado por el *Google News Lab* sobre los juegos olímpicos. A través de su labo-

ratorio, Simon Rogers junto a Christoforos Anagnostopoulos (2016) (*London's Imperial College* y *Mentat*) ofrecen una lectura interactiva del medallero olímpico, alternativa a la que habitualmente muestra el resto de medios. Concretamente, recalculan todos los datos del medallero normalizándolos, para poder realizar una comparación más justa entre los países a partir de los siguientes criterios de cada uno de los países: tamaño de su población, su PIB (ambos datos extraídos del *Banco Mundial*) y las búsquedas nacionales realizadas en *Google* (sobre los Juegos Olímpicos de 2016, sobre la participación deportiva y sobre la alimentación sana) (figura 6). De esta manera, si elegimos el criterio de la población, Bahamas pasa de la posición 51 de 87 en el medallero oficial a la primera posición. Y, si elegimos en cambio la opción del PIB, podemos observar que el país que mejor rentabiliza su producto interior bruto es la República de Fiya, con 63 medallas, frente a la posición 54 de 87 que ocupa en el medallero estándar.

“La cadena Univisión invitó a periodistas de once medios españoles y latinoamericanos para diseñar un fact-checking (verificación de hechos) en tiempo real que se aplicaría al tercer debate (y al debate presidencial final) entre Hillary Clinton y Donald Trump”

Por último, y en la misma línea, el rotativo español *El confidencial* ofrece otro interesante medallero alternativo (Escudero, 2016). En este caso, los factores normalizadores son las medallas por cada 100 deportistas, el número de medallas de oro por cada 100 deportistas y el número de deportistas por cada 100.000 habitantes (figura 7). En el caso de las medallas por cada 100 deportistas, Azerbaiyán se coloca en primera posición, seguido por Corea del Norte.

Según el número de medallas de oro por cada 100 deportistas, es Tayikistán, seguido de Kosovo, el que ocupa el primer lugar del ranking. Por último, si elegimos el número de deportistas por cada 100.000 habitantes, es la Mancomunidad de las Bahamas, seguida de la isla de Granada, quien manda en el medallero.

En el caso de las pasadas elecciones presidenciales norteamericanas, la propuesta de visualizaciones ha sido también ingente. La importancia intrínseca de las mismas y los resultados sorprendentes obtenidos han animado, todavía más si cabe, la oferta de explicaciones visuales del evento electoral. Si nos centramos exclusivamente en los aspectos visuales de esa oferta, encontramos dos propues-

tas especialmente interesantes.

La primera, publicada por *The Washington post* (Muyskens, 2016), denuncia con todo lujo de detalles cómo una parte importante de las gráficas utilizadas por Donald Trump en su campaña a través de las redes sociales están sesgadas, y no siempre a su favor (figura 8). Los sesgos o errores intencionados recurren a la batería típica de manipulaciones que suelen realizar sus poco éticos y escasamente profesionales *community managers* y diseñadores gráficos. Entre estos intentos de manipulación destaca, por encima de todos y para exagerar interesadamente la diferencia de los hechos, el sesgo de variación en las escalas, la no coincidencia del comienzo del eje de ordenadas (eje Y) en el valor 0 en el caso de las gráficas de columnas, o el de las abscisas (eje X) en el caso de las gráficas de barras.

La segunda visualización sobre esas elecciones que queremos destacar la propone la cadena *Univisión*. Esta cadena invitó a una serie de periodistas de once medios españoles y latinoamericanos para diseñar un *fact-checking* (verificación de hechos) en tiempo real que se aplicaría al tercer debate (y también al debate presiden-

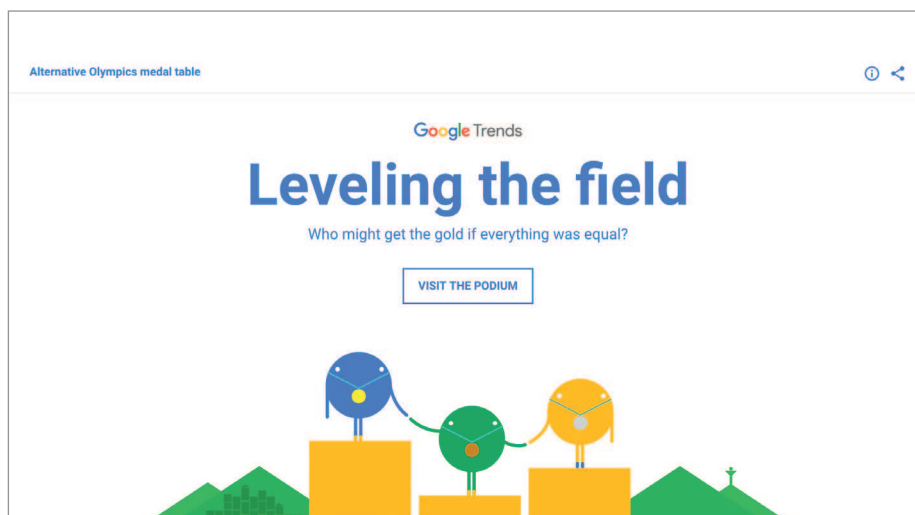


Figura 6. Medallero olímpico alternativo de Google News Lab
<https://landing.google.com/altmedaltable/results>



Figura 7. Medallero olímpico alternativo de *El confidencial*
<https://goo.gl/PTJQMI>



Figura 8. Artículo de *The Washington post* sobre la manipulación de las gráficas en la campaña de Trump
<https://www.washingtonpost.com/graphics/politics/2016-election/trump-charts>

cial final) entre Hillary Clinton y Donald Trump (figura 9) (Gómez; Liendo; Melgar, 2016). El resultado es una interesante visualización en la que cada una de las participaciones de los candidatos en el debate se representa como una superficie rectangular y en el que se colorean en rojo cada uno de los hechos falsos o menti-

ras incluidas en la intervención y en naranja los comentarios engañosos, además de la duración de los mismos respecto al tiempo total de la intervención. Si se sobrepasan las áreas coloreadas en rojo o naranja con el ratón podemos descubrir las declaraciones falsas o dudosas de cada uno de los candidatos, y si las clicamos obtenemos un comentario que explica por qué son considera-

Register for free at <https://www.scipedia.com> to download the version without the watermark

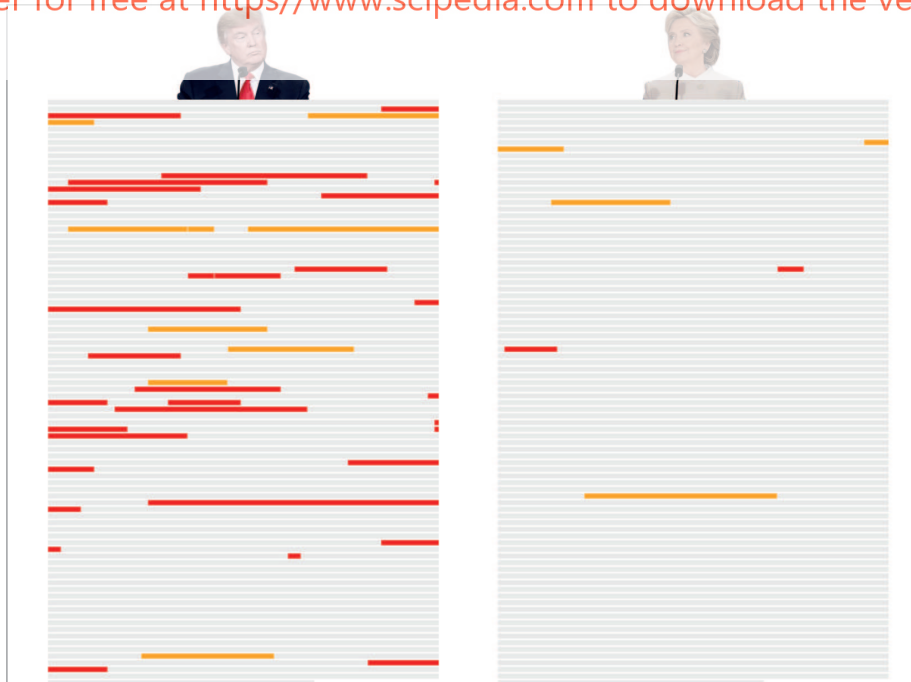


Figura 9. Visualización interactiva de los hechos falsos y dudosos incluidos en el tercer debate y en el debate presidencial final entre Hillary Clinton y Donald Trump
<http://www.univision.com/noticias/elecciones-2016/interactivo-todas-las-mentiras-de-trump-y-clinton-frente-a-frente-y-en-un-vistazo>

comentario que explica por qué son considera-

Dejando a un lado los hechos noticiables de los juegos olímpicos y las elecciones norteamericanas, la siguiente visualización que queremos destacar está relacionada con los contrastes existentes entre las diferentes culturas (McCandles, 2016). Concretamente, se intenta visualizar cada uno de los colores que se asocian con un concepto concreto dependiendo de la cultura a la que se pertenezca. O, dicho de otra manera, el significado de un color en cada cultura, Para ello se construye una matriz circular. En uno de sus ejes (el que determina la latitud en la matriz

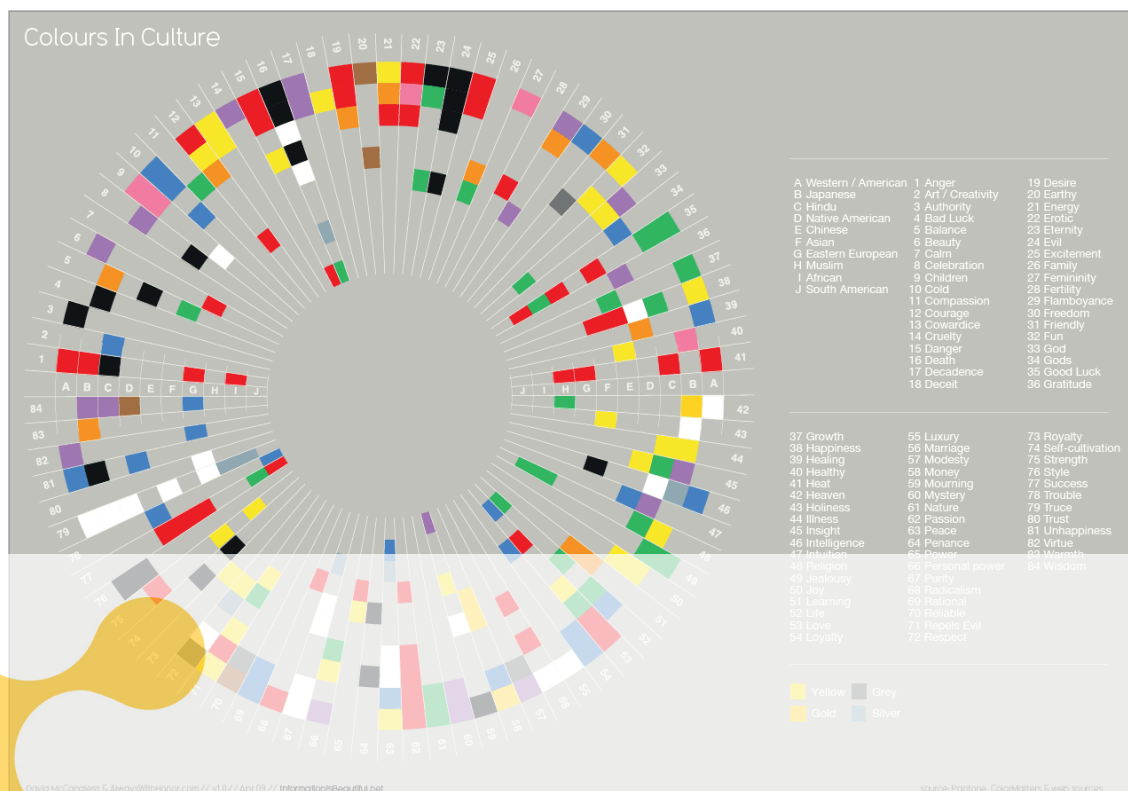


Figura 10. Visualización del significado de un color en cada cultura
<https://goo.gl/S2Oe2a>

SCIPEDIA

circular) se representan mediante letras mayúsculas las diferentes culturas existentes (de la A [cultura occidental o americana], a la J [cultura sudamericana]) (figura 10). En el otro eje (el que determina la longitud) se representan mediante un número cada uno de los conceptos representados, desde el 1 (enfado) al 84 (sabiduría). Las celdas creadas por la intersección de un concepto y una cultura se rellenan con el color (si existe) que se asocia con ese concepto concreto en esa cultura determinada (figura 10). Sorprende comprobar que el blanco está asociado con el concepto de tregua en todas las culturas menos en la cultura de los nativos americanos y en la de los asiáticos. Esta visualización ha sido utilizada como ilustración de la portada del libro *Information is beautiful* de David McCandless publicado en 2016 en el Reino Unido.

Queremos finalizar esta selección de visualizaciones interesantes por el conjunto de datos representado con una que se acerca al mundo de la filosofía. En este caso, desde la biblioteca de la *Universty of Iowa* (2016), se lanza una propuesta visual que intenta representar los aforismos (y sus dependencias lógicas) que conforman una de las obras, el *Tractatus Logico-Philosophicus* (1921), del filósofo más importante del s. XX, el austriaco Ludwig Wittgenstein (figura 11). Cada aforismo se representa con un punto. Si se clica en el punto,

podemos elegir entre la versión alemana del aforismo o las traducciones al inglés de Ogden o de McGuinness. Las dependencias lógicas entre los aforismos se representan mediante una estructura en forma de árbol. Hay dos botones de colores que los conectan. El botón verde en la esquina superior izquierda del mapa conduce a una representación similar del *Prototractatus*, una versión temprana del *Tractatus*.

“La mayoría de los medios de comunicación han creado infinidad de propuestas visuales para completar sus textos y poder comunicar de forma efectiva a sus lectores y espectadores los aspectos relevantes contenidos en esos conjuntos de datos”

4. Tecnología

El ecosistema de propuestas tecnológicas para la creación de visualizaciones es muy extenso. Cada semestre podemos comprobar como aparecen nuevas herramientas y cómo otras caen en desuso hasta su desaparición. En este contexto, para cubrir sus necesidades de visualización,

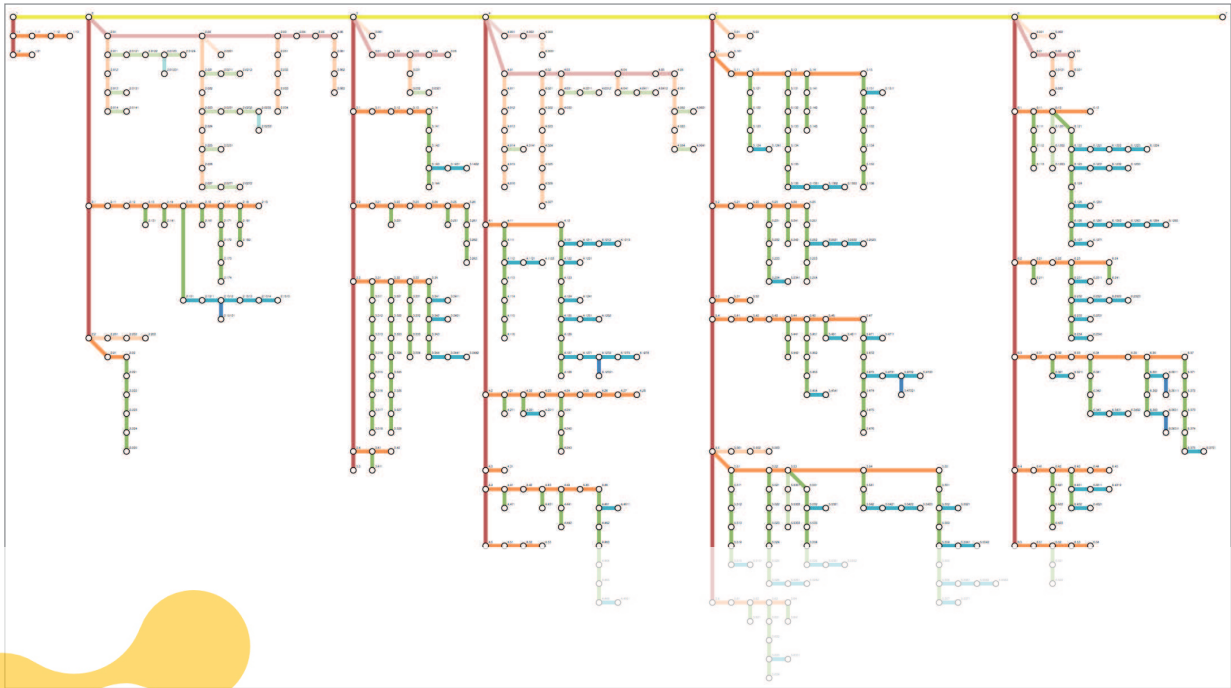


Figura 11. Propuesta visual de la obra *Tractatus logico-philosophicus* (1921) de Ludwig Wittgenstein realizado por la biblioteca de la University of Iowa
<http://tractatus.lib.uiowa.edu/tlp>

SCIPEDIA

muchas compañías adoptan tecnologías comerciales contrastadas como, por ejemplo, *Tableau*, *Qlik* o *MicroStrategy*. Pero otras, en cambio, apuestan por desarrollar sus propias propuestas visuales en HTML5, JavaScript, Python o R.

De todas formas, sin perder de vista este escenario, es posible realizar una pequeña selección de algunas de las noticias tecnológicas más llamativas de 2016 relacionadas con la representación de información.

“Para cubrir sus necesidades de visualización, muchas compañías adoptan tecnologías comerciales como *Tableau*, *Qlik* o *MicroStrategy*. Otras apuestan por desarrollar sus propias propuestas visuales en HTML5, JavaScript, Python o R”

Comencemos, primero, por el apartado de las ausencias. Entre éstas destaca la desaparición definitiva de la herramienta *Many Eyes* de IBM. Lanzada en 2007, esta propuesta supuso una verdadera democratización en el campo de la visualización al permitir que casi cualquier persona pudiese cargar, visualizar (incluso con representaciones interactivas) y compartir fácilmente y de forma pública sus datos en línea, sin tener que

instalarse ninguna aplicación. La clave fue permitir, hace ya casi una década, ejecutar novedosamente todas estas operaciones con muy pocos clics. En la segunda mitad de 2015 se clausuró la herramienta. En 2016, se cerró definitivamente la web con los datos y las visualizaciones creadas. La apuesta tecnológica actual de IBM para la gestión y visualización de datos se concentra alrededor de la herramienta *Watson Analytics*.

<https://www.ibm.com/analytics/watson-analytics/us-en>

Por otro lado, pasemos ahora a hablar brevemente sobre R.

<https://www.r-project.org>

R es un lenguaje y un entorno de programación para el análisis estadístico y la creación de gráficos. Es un proyecto GNU desarrollado en los *Bell Laboratories* (anteriormente AT&T, ahora *Lucent Technologies*) por el equipo de John Chambers. Además, se encuentra disponible como *software libre* bajo los términos de la *GNU General Public License* de la *Free Software Foundation*.

En muchos casos, R se considera como un estándar tecnológico para el tratamiento integral de los datos. Sin embargo, hay que señalar la sucesiva aparición de diversos paquetes que mejoran sus funciones en la tarea de la representación gráfica. Entre éstas podemos destacar de forma especial la consolidación durante 2016 de:

- *ggplot*
<http://ggplot.yhathq.com>

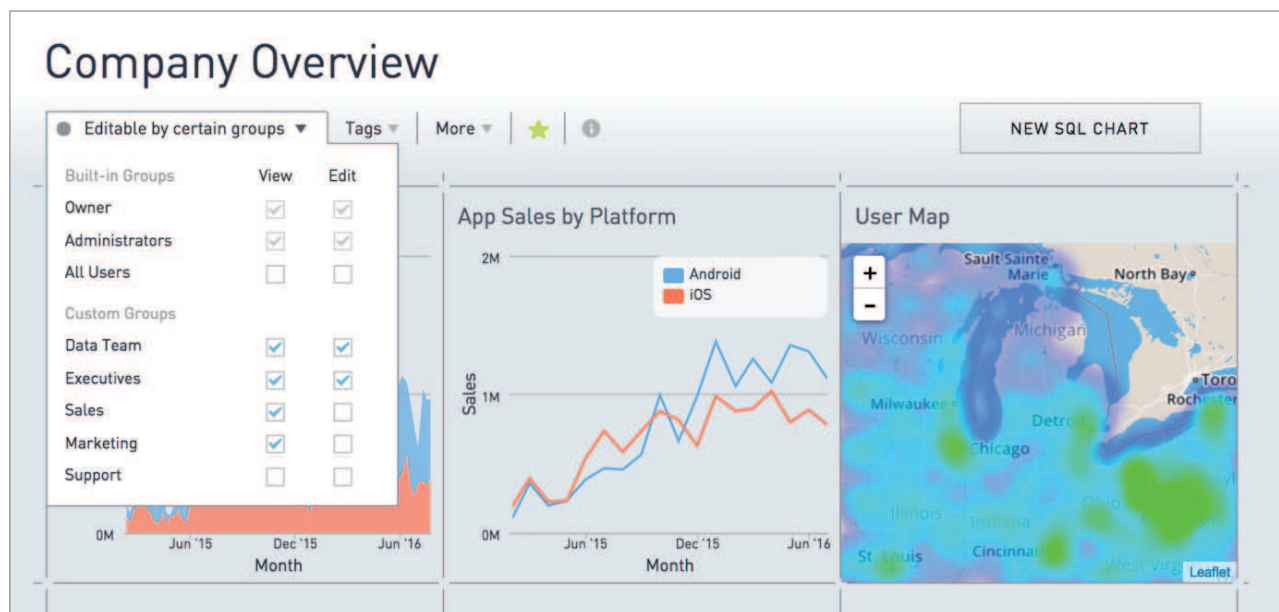


Figura 12. Propuesta parcial de cuadro de mando realizado con *Periscope Data*
<https://www.periscopedata.com/?src=logo>

- *dplyr*
<https://cran.r-project.org/web/packages/dplyr/index.html>
- *ggvis*
<http://ggvis.rstudio.com>

La ventaja de todas estas herramientas reside en que ofrecen un sistema de representación para *R* que permite gestionar grandes conjuntos de datos, interrogar directamente una base de datos y crear visualizaciones interactivas para la web dotadas de un alto estándar profesional con la contrapartida de una utilización limitada de código. El resultado refuerza la opción de *R* como alternativa sólida para la visualización de información, sobre todo, para el colectivo de los desarrolladores y las empresas que prefieren no utilizar otras alternativas comerciales.

En tercer lugar, queremos resaltar también la propuesta de nuevas herramientas para la creación de *dashboards*. Un *dashboard* (o cuadro de mando) es un dispositivo visual (una forma de visualización, no una tecnología concreta) que ofrece la información necesaria, crítica e importante para cubrir uno o más objetivos por parte de un usuario. En este sentido, representa la información de forma visual, consolidada y sistematizada en una única pantalla para que pueda ser monitorizada a voluntad. Dentro de este tipo de herramientas queremos destacar las que nos ofrecen desde *Periscope Data* y desde *Airbnb*.

Periscope Data propone una herramienta que permite gestionar e interrogar nuestras bases de datos. También puede visualizar los resultados de cualquier consulta con gráficos integrados (grá-

fica, tabla, mapa o diagrama de dispersión) que admiten un gran abanico de opciones de personalización. Esos resultados pueden ser compartidos con un número ilimitado de usuarios de forma instantánea mediante un sistema de gestión de permisos. Estos procesos se complementan con informes enviados por correo electrónico a esos usuarios (figura 12).

<https://www.periscopedata.com/?src=logo>

“Un *dashboard* (o cuadro de mando) es un dispositivo visual que ofrece la información necesaria, crítica e importante para cubrir uno o más objetivos por parte de un usuario. Representa la información de forma visual, consolidada y sistematizada en una única pantalla para que pueda ser monitorizada a voluntad”

Airbnb también nos ofrece una herramienta, *Superset*, pero en este caso de código abierto, para la creación de cuadros de mando personalizados. Al igual que en el caso anterior, proporciona una forma rápida e intuitiva de crear y compartir de manera personalizable paneles visuales interactivos (figura 13).

<https://github.com/airbnb/superset>

Por último, aunque no sea una propuesta tecnológica en un sentido estricto, queremos destacar también un nuevo modelo de negocio

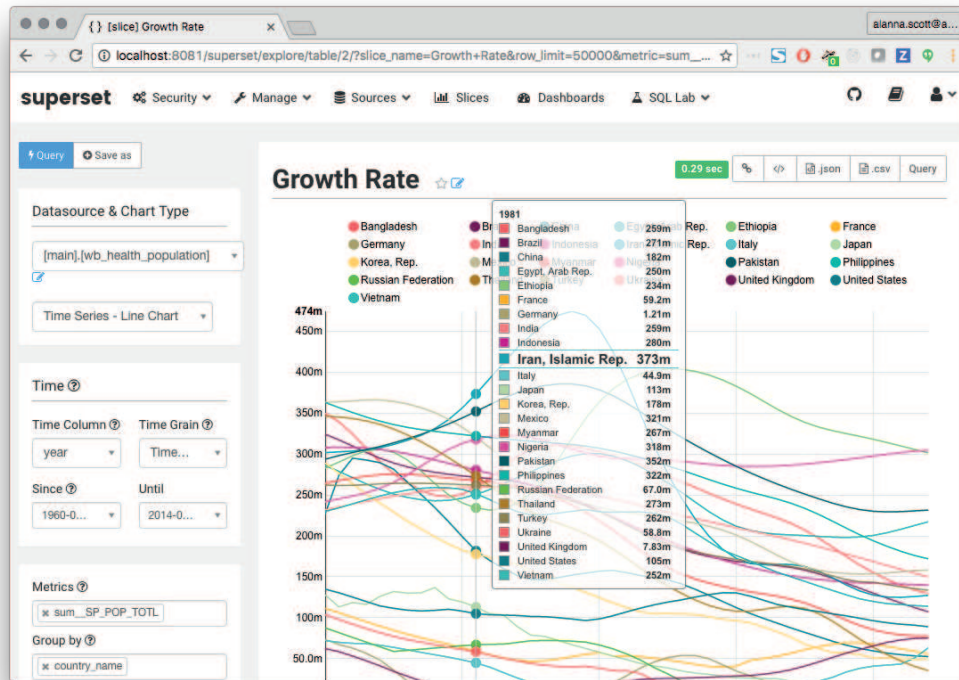


Figura 13. Propuesta parcial de cuadro de mando realizado con *Superset* de *Airbnb*
<https://github.com/airbnb/superset>

aparecido alrededor de los datos y su visualización en el contexto de las empresas periodísticas.

ProPublica es un medio de información digital independiente que produce periodismo de investigación de interés público. En esta publicación, el uso intensivo de datos y su visualización interactiva son piezas críticas que articulan su proyecto informativo. En febrero de 2014, lanzaron el proyecto experimental *ProPublica Data Store*.
<https://www.propublica.org/datastore>

Este proyecto nace como el repositorio donde almacenar los datos involucrados en su praxis periodística. Dos años más tarde, se ha abandonado la dimensión experimental y se consolida como un verdadero modelo de negocio. La idea que subyace es bien sencilla. Los lectores del medio puedan ceder de forma documentada sus datos si son susceptibles de poder integrarse en un proceso de investigación periodística. Esta cesión puede ser remunerada por parte del medio. Por otro lado, junto a los derivados de su propia praxis periodística, la web *ProPublica Data Store* ofrece esos conjuntos de datos cedidos para aquellos que quieran explotarlos. Estos datos se ofrecen ya refinados, procesados, enriquecidos y completados. Algunos son gratuitos, pero la mayoría son de pago. En el caso del pago, los precios pueden variar si la explotación es comercial, académica

o periodística. Además, las agencias de noticias pueden enviar conjuntos de datos que han sido refinados, analizados, documentados y utilizados en informes ya publicados. *ProPublica* se encarga de administrar su comercialización en nombre de esas organizaciones. El beneficio del proyecto fue de 30.000 dólares en los primeros cinco meses y ya ha acumulado más de 200.000 dólares desde el lanzamiento. Los ingresos de este último año han superado la suma de los dos últimos años.

No queremos finalizar este informe sin destacar también, aunque no pueda clasificarse en ninguno de los tres apartados anteriores, una futura novedad editorial relacionada con la disciplina de la visualización de información. Existen infinidad de monografías que pueden ayudarnos a introducirnos en el mundo de la visualización de información. Algunas, con una marcada dimensión conceptual; otras, con una clara orientación práctica.

Todas ellas, sin duda, contribuyen a la alfabetización visual de sus lectores. Sin embargo, muy pocas están orientada a facilitar este tipo de alfabetización en el público infantil. Para cubrir ese espacio vacío, **Liv Buli** (periodista de datos) y **Abigail Ricarte** (diseñadora) (2016) deciden escribir *Vizkidz*, una serie de libros sobre visualización de datos para niños. En estos libros, a través de una serie de personajes (Penelope Pie,

Laney Line, Barnaby Bar y Bertie Boxplot, todos con apellidos que hacen referencia a tipos de gráficas) y sus aventuras tratan de explicar los conceptos básicos que sustentan la visualización de datos y que permiten interpretar, analizar y crear representaciones visuales comunicacionalmente eficientes. Para conseguir los fondos necesarios para llevar adelante la producción, publicación, promoción y distribución de este proyecto editorial, las autoras han iniciado una campaña de *crowdfunding* (financiación colectiva) en la plataforma de *Kickstarter*.

<https://www.kickstarter.com/projects/414625356/vizkidz-a-series-of-books-on-data-visualization-fo>

5. Referencias

Aisch, Gregor; Buchanan, Larry (2016). "A visual history of which countries have dominated the summer olympics". *The New York times*, 22 agosto. http://www.nytimes.com/interactive/2016/08/08/sports/olympics/history-olympic-dominance-charts.html?_r=1

Buli, Liv; Ricarte, Abigail (2016). "Vizkidz: A series of books on data visualization for kids". *Kickstarter*. <https://www.kickstarter.com/projects/414625356/vizkidz-a-series-of-books-on-data-visualization-fo>

Casselman, Ben; Conlen, Matthew; Fischer-Baum, Reuben (2016). "Gun deaths in America". *FiveThirtyEight*. <http://fivethirtyeight.com/features/gun-deaths>

Escudero, Jesús (2016). "El medallero alternativo de los Juegos de Río". *El confidencial*, 22 agosto. <https://goo.gl/2VEChM>

Gómez, Juanje; Liendo, Olivia; Melgar, Luis (2016). "Interactivo: todas las mentiras de Trump y Clinton frente a frente y en un vistazo". *Univisión*, 19 octubre.

<http://www.univision.com/noticias/elecciones-2016/interactivo-todas-las-mentiras-de-trump-y-clinton-frente-a-frente-y-en-un-vistazo>

Lupi, Giorgia; Posavec, Stefanie (2016). *Dear data*. <http://www.dear-data.com>

McCandless, David (2016). "Colours in culture". *Information is Beautiful*. <https://goo.gl/jZytCP>

Muyskens, John (2016). "Most of Trump's charts skew the data. And not always in his favor". *The Washington post*, 31 octubre. <https://www.washingtonpost.com/graphics/politics/2016-election/trump-charts/>

Posavec, Stefanie (2016). "(En)tangled Word Bank". Stefanie Posavec. <http://www.stefanieposavec.co.uk/entangled-word-bank>

Rogers, Simon; Anagnostopoulos, Christoforos (2016). "Leveling the field. Who might get the gold if everything was equal?". *Google Trends*. <https://landing.google.com/altmedatable/>

Toronto Symphony Orchestra (2016). "TSO visual listening guide". *Toronto Symphony Orchestra*. <https://www.tso.ca/listening-guides>

Universty of Iowa Library (2016). "University of Iowa Tractatus Map". *Universty of Iowa*. <http://tractatus.lib.uiowa.edu/tlp>

Mario Pérez-Montoro

Universitat de Barcelona

Departamento de Biblioteconomía, Documentación y Comunicación Audiovisual
perez-montoro@ub.edu

Si te interesan los

INDICADORES EN CIENCIA Y TECNOLOGÍA,

y todos los temas relacionados con la medición de la ciencia, tales como:

Análisis de citas, Normalización de nombres e instituciones, Impacto de la ciencia en la sociedad, Indicadores, Sociología de la ciencia, Política científica, Comunicación de la ciencia, Revistas, Bases de datos, Índices de impacto, Políticas de open access, Análisis de la nueva economía, Mujer y ciencia, etc.

Entonces **INCYT** es tu lista. Suscríbete en:

<http://www.rediris.es/list/info/incyt.html>